

DOCUMENT RESUME

ED 294 897

TM 011 533

AUTHOR Reynolds, Arthur J.; Bezruczko, Nikolaus
TITLE Assessing the Construct Validity of a Life Skills Competency Test.
PUB DATE Apr 88
NOTE 2lp.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 6-8, 1988).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Tests; *Construct Validity; *Daily Living Skills; Elementary School Students; Grade 8; Junior High Schools; Mathematics Tests; *Minimum Competency Testing; Multiple Choice Tests; Reading Tests; Standardized Tests; Urban Environment
IDENTIFIERS *Chicago Minimum Proficiency Skills Test

ABSTRACT

The proliferation of minimum competency testing in recent years has not resulted in a critical analysis of the construct validity of such tests. The present study examined the dimensionality and construct validity of an urban life skills competency test (the 1987 eighth-grade Chicago Minimum Proficiency Skills Test (MPST) via linear analysis. The MPST is a 63-item multiple choice test. Three computerized random samples of 880, 879, and 879 Chicago public elementary school students (Grade 8) were selected for analysis. Results of a double cross-validation factor analysis identified one factor underlying performance on the test. This factor was stable across three independent samples. This derived factor was highly correlated with reading and math sub-tests of the Iowa Test of Basic Skills. Thus, life skills competency tests may not be operationally different than traditional achievement tests. Additional validity studies are needed. Four tables are included. (Author/TJH)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED294897

Life Skills Competency

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

ARTHUR J. REYNOLDS

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Assessing the Construct Validity of a Life Skills Competency Test

Arthur J. Reynolds and Nikolaus Bezruczko
Chicago Public Schools

Running Head: LIFE SKILLS COMPETENCY

A paper presented at the 1988 Annual Meeting of the National
Council for Measurement in Education, New Orleans, April,
1988.

TM 011 533

Abstract

The proliferation of minimum competency testing in recent years has not resulted in a critical analysis of the construct validity of such tests. The present study examined the dimensionality and construct validity of an urban life skills competency test. Results of a double cross-validation factor analysis identified one factor underlying performance on the test. This factor was stable across three independent samples. This derived factor was highly correlated with reading and math subtests of the ITBS. Thus, life skills competency may not be operationally different than traditional achievement tests. Additional validity studies are needed.

This paper explores the construct validity of an instrument measuring minimum competency life skills. Although minimum competency tests (MCT) have been mandated by almost all state governments, few researchers agree on its conceptual foundation (Haney & Madaus, 1978). Many investigators do not believe that MCTs differ in nature from standardized achievement tests. This study performs a double cross-validation factor analysis of the 1987 eighth grade Chicago Minimum Proficiency Skills Test (MPST). Three related questions are investigated: (1) What is (are) the dimension(s) of life skills MCTs? (2) Are the empirically derived factors of MCTs stable over different sample populations? (3) Are the underlying factors associated with traditional achievement tests.

Research Perspective

Although MCTs continue to be implemented in school districts and states across the country, there is no uniform definition or common conceptual framework for competency (Haney & Madaus, 1978). While there is little disagreement over the definition of competency (i.e., necessary skills and/or abilities to function adequately in the world), the operationalization of this meaning has produced a variety of competency tests. For example, there are minimum achievement tests, life skills tests, academic readiness tests, and reading & math MCTs. The degree to which these tests measure

common underlying dimensions is unclear as is the meaning of derived scores.

The label "minimum competency" begs the question similar to that of intelligence: "Minimum competency for what?" It is assumed that competence, however defined, resides in the individual and can be generalized across situations. LaDuca (1980) argues that competence is not a general ability but is situation-specific in regard to some specified criterion behavior. Even in a highly specific area such as medicine there is little consensus on the dimensions of clinical competence (Forsythe, McGaghie, & Friedman, 1986). This suggests the magnitude of the problem in determining and interpreting competence in a more general area of behavior such as education.

This concern is related to the general issue of construct validity. Although Cronbach (1971) has emphasized construct validity as the basis on which to establish test meaning and interpretability, it is often ignored in test validation in favor of content and predictive considerations. Hambleton, Swaminathan, Algina, and Carlson (1978) noted this lack of construct validity for MCTs and other criterion-referenced tests. They suggested more work in this area such as exploratory and confirmatory factor analysis. Jaeger and Tittle (1980) also indicated the need for validation of competency tests. With the exception of some related studies

(i.e., Forsythe et al., 1986), little work has been done in this regard.

Another issue not addressed in the competency research is the relationship between basic achievement tests and MCTs. If one posits, as many do, that the goal of competency programs is the "identification of minimum abilities each student must attain in order to function in society" (cf. Haney & Madaus, p. 463), it would be expected that competency can be differentiated from achievement in school subjects. While McPhee and Kerr (1985) studied the relationship between achievement and curriculum competency, research on life skills competency has not been done. Investigating the differential validity of life skills competency tests may then help shed light on their construct validity.

In this paper, the construct validity and dimensionality of a life skills competency test is analyzed through linear factor analysis. Factor analysis is a recommended technique for assessing dimensionality, especially when there have been no prior applications (Carmines & Zeller, 1979; Kim & Mueller, 1978). However, Thorndike (1978) advocated cross-validation of factor analytic results because of the threat of sample-specific covariation. Most studies using factor analysis fail to address this cross-validation concern. This recommendation will be used in the present study. It was hypothesized that there is no dominant underlying factor in

the MPST and that life skills competency can be differentiated from basic achievement.

Method

Subjects

Three computerized random samples of 880, 879, and 879 Chicago public elementary students were selected for the study. These samples represent about 10% of over 26,000 eighth grade students taking the MPST in spring 1987. They were taking the MPST for the first time. Gender and racial characteristics of sample 1 indicated that 51% were male, 49% female, and 17%, 65%, and 18% were white, black, and hispanic.

Measures

MPST. The Chicago test is a 63-item, multiple choice test designed to measure minimum proficiency in consumer life skills. It is divided into three general areas: (1) Language Arts (LGA), (2) Computation (CPT), and (3) Problem Solving (PBS). Language Arts items are reading and vocabulary-oriented. Computation items are composed of simple math and arithmetic operations. Problem Solving items are reading-oriented but focus on solving or addressing a particular situation (i.e., who to call or where to go in an emergency). Each of the three general components of the test are further separated into seven consumer components, resulting in a 21-subtest matrix.

These seven content-specific variables include areas in which a person should be expected to be minimally proficient after high school including (a) Finance, (b) Community resources (i.e., housing, parks), (c) Communication (i.e., information from the print media), (d) Health, (e) Occupation (i.e., information about jobs), (f) Government (i.e., knowledge of government offices), and (g) Transportation (i.e., travel). The test requires students to apply reading and math skills to everyday life events such as map reading, deciphering warning labels, balancing a checkbook, interpreting reading selections, and identifying basic information about government and health. For example, item 22 (LGA Finance) reads "Mary wishes to buy a new car and she needs a loan from the bank to buy the car. What is a 'loan'?" The alpha reliability of the test for sample 1 was .88. Passing the MPST is a high school graduation requirement.

Iowa Test of Basic Skills (ITBS). To examine the construct validity and dimensionality of the MPST, the derived construct(s) were be correlated with reading and math subtests of the spring 1987 ITBS (Level 14 Form 7/8).

Data Analysis

To test the underlying structure of the test, a variety of exploratory factor analytic extraction methods were employed including principal components, generalized least squares, and maximum likelihood analyses. In this process,

the 21 subtests of the MPST were used as the unit of analysis. Each subtest included from one to four dichotomously coded items. Although the small number of items for each subtest restricts the range of resulting correlations, subtest analysis appears preferable to the analysis of dichotomous items. While three of the subtests include only one item, the use of such variables is supported in exploratory analysis (Kim & Mueller, 1978). To prevent subtests with an unproportionally high number of items to unduly influence factor loadings, six items were removed from three subscales before the analysis.

Independent factor analyses were performed on the three samples and their factor loadings compared. This method of cross-validation is consistent with more specific procedures such as Procrustes rotation and confirmatory factor analysis (Thorndike, 1978). The criteria used for selecting the number of factors to retain were empirical and interpretive. Factors having eigen values greater than one (Kaiser criterion) were initially selected for inclusion. However, only those that could be clearly interpreted were retained.

Results

Subtest means and standard deviations for the validation sample (sample 1) are reported in Table 1. The relatively large standard deviations for the small number of items used indicated the diverse range of scores obtained by students.

Subtest intercorrelations also indicated moderate variance among the test variables as values ranged from .07 to .50.

Insert Table 1 about here

Table 2 presents the factor loadings from the rotated two-factor solutions of the validation sample and the two concurrent cross-validations samples (samples 2 and 3). With regard to the validation sample (sample 1), factor 1 explained 30% of the variance in subtest scores and was most associated with computation variables as well as two problem solving scales. It was interpreted to be a math/computation factor. Nine subtests loaded highly on this numerical component. CPT Finance and CPT Transportation had the highest loadings. Factor 2 appears to be a language/problem solving factor, although it explains only 5.5% of the variance in scores.

Insert Table 2 about here

However, as further displayed in Table 2, the two-factor model did not hold up over two replications. While sample 2 was nearly identical to the validation sample in the percent of variance explained, the magnitude of factor loadings, and the factor interpretations, the pattern of factor loadings in sample 3 was the reverse of the first two samples.

Language/problem solving became the dominant factor and computation was subsidiary, although the amount of variance explained and the magnitude of factor loadings were similar to the first two samples.

This contradictory result for sample 3 indicates that factor 2 may not be independent of factor 1; in fact, it may be an artifact of it. Although the two-factor structure satisfied the criteria for factor selection (i.e., eigen values greater than one and clear interpretation), it was not stable over cross-validation.

Table 3 presents the cross-validation results of a revised one-factor model. As shown, the derived factor is relatively stable across samples. Both the proportion of variance explained and the magnitude of factor loadings is consistent. For example, loadings on CPT Communication and CPT Transportation are consistently high across samples (over .60) as are most other variables.

Insert Table 3 about here

There were some differences in the composition of the factor across samples. Language Arts subtests (i.e., government, occupation, communication) were more prominent in sample 1 and 2. In addition, the high loading of CPT Finance (.53) was unique to sample 1.

Thus, given the representation of many general and specific content areas on the derived factor, a general competency construct is suggested, although a large proportion of variance (70%) remained unaccounted for.

Correlation with Achievement Measures

To further examine the above results and investigate the construct validity of the MPST, the one-factor model was correlated with 1987 ITBS subtests. Factor scores were computed by multiplying each factor loading of .50 or above by individual students' subtest scores and then summing the linear combinations. As Table 4 indicates, the underlying competency construct is highly correlated with both reading comprehension and total math subtest scores (above .70 for all samples), suggesting a general achievement construct. Math computation was less correlated with the derived factor.

Insert Table 4 about here

Thus, it appears the underlying competency factor has strong components of both reading and math achievement. Approximately one-half of the variance in achievement test scores can be explained by this underlying competency factor. The prediction of ITBS scores from competency is by no means perfect. However, the strong relationship between the two sets of variables indicate the difficulty of separating achievement from life competency.

Discussion

The results of this study indicate that (1) there is one relatively strong underlying dimension to the life skills MPST, (2) the factor is stable across three independent samples of eighth grade public school students, and (3) it is highly correlated with concurrent ITBS reading and math subtest scores.

The rationale for designing and implementing MCT programs is that competency tests, especially life skills competency tests, give an indication of how well one is prepared to conduct their affairs in the real world after leaving high school. While there were other reasons for instituting such programs (i.e., accountability and perceptions of declining achievement (Ross, 1982), the goal and measurement of competency tests was intended to be differentiated from basic achievement tests. This study indicates a measure of life skills competency is not dissimilar in meaning to achievement tests as measured by the ITBS. The correlations between the competency factor and ITBS reading and math subtests (.70 and above) are quite similar to the intercorrelations of ITBS subtests (i.e., reading - math total $r = .73$). In addition, ITBS test-retest correlations are also surprisingly consistent with the competency-ITBS relationship. These results suggest that competency and achievement are more similar than different.

Moreover, the observed correlation between the competency and achievement constructs is consistent with the recommended reliability (i.e., correlation between two similar measures) of a test (i.e., .70 or above). The primary underlying dimension of competency appears to be highly similar to math and reading achievement, thus, the MPST may best be seen as a measure of academic performance rather than life competency per se.

An additional finding of the study is the importance of cross-validation in factor analysis. Without the double cross-validation employed, a two-factor model would have been mistakenly adopted in which a one-factor model was clearly more desirable. Cross-validation should be an essential feature of all factor analyses as well as other multivariate techniques.

A limitation of the present study which may have restricted results is the small number of variables used to define a subtest. Some scales were no better than dichotomous items while others had up to four items. The lack of interval scale variables can increase nonrandom measurement error and probably hindered the model from extracting additional factors. This measurement error may have also been a partial explanation of why the two-factor model was unstable and why other minor factors were unspecifiable.

A goal of the study was to clarify the underlying structure of life skills competency. The findings suggest that scores from a life skills MPST are highly comparable to those of standardized achievement tests. Thus, competency scores may not be an indicator of competency as much as they are an index of achievement. Snow (1986) stressed the need for research on the generalization of MCT scores. Present findings suggest that such scores are generalizable to standardized achievement tests.

Future Directions

While the present study indicates the similarity of life skills competency and basic achievement, additional research is needed. Investigating the predictive validity of life skills competency may help clarify the nature of the construct. A comparison of the predictive validity (i.e., on job success) of life skills competency and basic achievement would also provide further information on the similarity of these tests. Finally, a theory of competency is needed to help guide empirical work on competency and to clearly distinguish competency from achievement.

References

- Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Quantitative applications in the social sciences (no. 6-017). Beverly Hills: Sage.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), Educational Measurement (2nd Ed.). Washington: American Council on Education.
- Forsythe, G. B., McGaghie, W. C., & Friedman, C. P. (1986). Construct validity of medical competence measures: A multitrait-multimethod matrix study using confirmatory factor analysis. American Educational Research Journal, 23, 315-326.
- Hambleton, R. K., Swaminathan, H., Algina, J., and Coulson, D. B. (1978). Criterion-referenced tests and measurement: A review of technical issues and developments. Review of Educational Research, 48, 1-47.
- Haney, W., & Madaus, G. (1978). Making sense of the competency testing movement. Harvard Educational Review, 17, 462-484.
- Jaeger, R. M., & Tittle, C. K. (Eds.). (1980). Minimum Competency Achievement Testing. Berkeley: McCutchan
- Kim, J. O., & Mueller, C. W. (1978). Factor analysis: Statistical methods and practical issues. Quantitative applications in the social sciences (no. 07-14). Beverly Hills: Sage.

- LaDuca, A. (1980). The structure of competence in health professions. Evaluation & the Health Professions, 3, 253-258.
- McPhee, S. A., & Kerr, M. E. (1985). Scholastic aptitude and achievement as predictors of performance on competency tests. Journal of Educational Research, 78, 186-190.
- Nunnally, J. C. (1978). Psychometric theory. New York: McGraw-Hill.
- Ross, D. D. (1982). Competency based education: Understanding a political movement. Educational Forum, 46, 483-490.
- Snow, R. E. (1986). Individual differences and the design of educational programs. American Psychologist, 41, 1029-1039.
- Thorndike, R. M. (1978). Correlational procedures for research. New York: Gardner Press.

Table 1

Subtest Means and Standard Deviations for Sample 1 (N = 880)

<u>Subtest</u>	<u>M</u>	<u>SD</u>
LGA Finance	.57	.50
LGA Resources	.68	.47
LGA Communication	1.71	.90
LGA Health	2.76	1.03
LGA Occupation	1.23	.68
LGA Government	2.78	.92
LGA Transportation	1.38	.65
CPT Finance	1.48	.88
CPT Resources	2.10	.85
CPT Communication	3.00	1.02
CPT Health	.92	.74
CPT Occupation	1.85	.83
CPT Government	2.13	.79
CPT Transportation	2.36	.83
PBS Finance	1.32	.71
PBS Resources	1.31	.70
PBS Communication	2.99	1.00
PBS Health	2.90	1.01
PBS Occupation	.80	.40
PBS Government	1.88	.96
PBS Transportation	2.32	.81
Total	42.93	10.10

Note. LGA = Language Arts, CPT = Computation, PBS = Problem Solving. Total test included 57 items.

Table 2

Loadings of a Two-Factor Principal Components Varimax
Solution for Validation and Cross Validation Samples

Variable	Original Sample		Cross Validation			
			Sample 2		Sample 3	
	1	2	1	2	1	2
LGA Finance						
LGA Resources		.45		.44	.45	
LGA Communication		.58		.46	.58	
LGA Health		.57		.50	.56	
LGA Occupation		.47		.53	.56	
LGA Government				.46	.45	
LGA Transportation				.44	.47	
CPT Finance	.48		.54			.54
CPT Resources	.69		.64			.61
CPT Communication	.62		.58			.60
CPT Health	.61		.46			.57
CPT Occupation	.52		.51			.54
CPT Government	.61		.56			.57
CPT Transportation	.69		.68			.69
PBS Finance	.48		.48			.47
PBS Resources		.59		.66	.56	
PBS Communication		.57		.56	.56	
PBS Health		.52		.49	.55	
PBS Occupation						
PBS Government		.64		.68	.57	
PBS Transportation	.61		.64			.58
R ²	29.6	5.5	27.1	6.0	27.3	6.5

Note. Only variables with factor loadings .44 or above are listed in the table. Original sample $N = 880$, sample 2 $N = 879$, sample 3 $N = 879$.

Table 3

Loadings of a One-Factor Principal Components
Solution for Validation and Cross Validation Samples

Variable	Original Sample	Cross Validation	
		Sample 2	Sample 3
	1	1	1
LGA Finance			
LGA Resources			
LGA Communication	.53	.54	.56
LGA Health	.52		.54
LGA Occupation		.50	
LGA Government	.51	.52	
LGA Transportation			
CPT Finance	.53		
CPT Resources	.62	.51	.55
CPT Communication	.71	.65	.63
CPT Health	.51		
CPT Occupation			
CPT Government	.66	.62	.59
CPT Transportation	.67	.65	.68
PBS Finance	.54	.52	.59
PBS Resources	.57	.53	.52
PBS Communication	.62		.61
PBS Health	.57	.56	.53
PBS Occupation			
PBS Government	.61	.63	.63
PBS Transportation	.67	.60	.60
R ²	29.6	27.1	27.3

Note. Only variables with factor loadings .50 or above are listed in the table.

Table 4

Correlations Between the Derived Competency Factor and ITBS Subtest Scores for Sample Groups

Derived Factor	ITBS Subtest		
	Reading Comprehension	Math Computation	Math Total
Sample 1 Factor	.72	.61	.77
Sample 2 Factor	.70	.57	.72
Sample 3 Factor	.70	.55	.72

Note. All correlations were significant at $p < .001$.